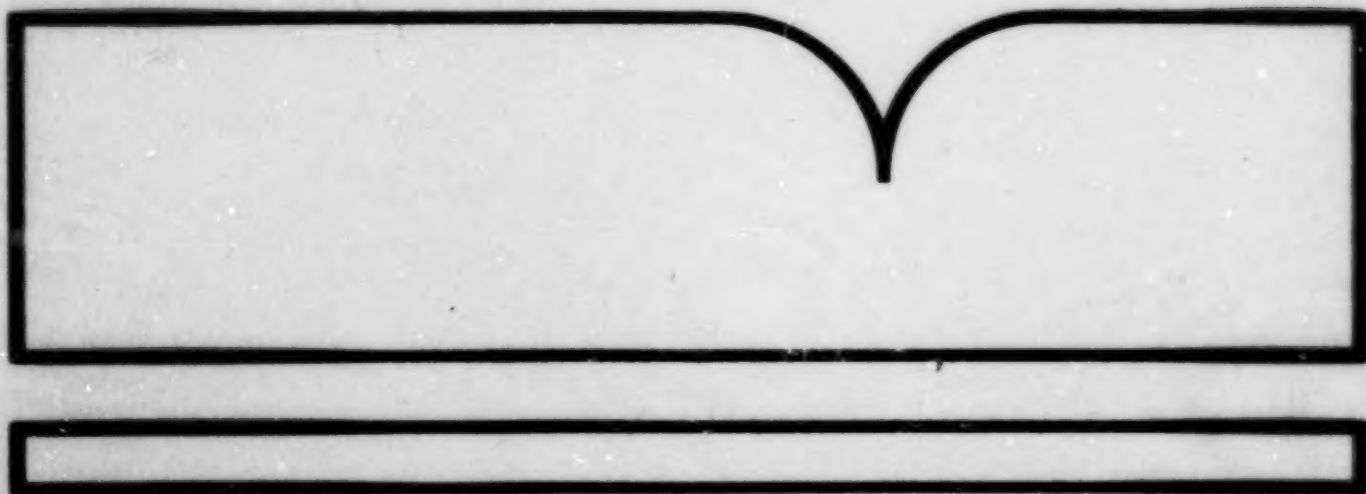


PB85-124675

Statistical Relationships between
Exposure and Dose

(U.S.) Environmental Monitoring Systems Lab.
Research Triangle Park, NC

Nov 84



U.S. Department of Commerce
National Technical Information Service

NTIS

EPA-600/D-84-274
November 1984

STATISTICAL RELATIONSHIPS BETWEEN EXPOSURE AND DOSE

by

David T. Mage
Environmental Monitoring Systems Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Cooperative Agreement CR808536-01-0

ENVIRONMENTAL MONITORING SYSTEMS LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

TECHNICAL REPORT DATA (Please read Instructions on the reverse before completing)		
1. REPORT NO. EPA-600/D-84-274	2.	3. RECIPIENT'S ACCESSION NO. PB80 124675
4. TITLE AND SUBTITLE Statistical Relationships Between Exposure and Dose	5. REPORT DATE November 1984	
	6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) David T. Mage	8. PERFORMING ORGANIZATION REPORT NO.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Environmental Protection Agency ERC Annex (MD-56) Research Triangle Park, N. C. 27711	10. PROGRAM ELEMENT NO.	
	11. CONTRACT/GRANT NO. CR808536-01-0	
12. SPONSORING AGENCY NAME AND ADDRESS ENVIRONMENTAL MONITORING SYSTEMS LAB. OFFICE OF RESEARCH AND DEVELOPMENT U.S. ENVIRONMENTAL PROTECTION AGENCY RESEARCH TRIANGLE PARK, NC 27711	13. TYPE OF REPORT AND PERIOD COVERED	
	14. SPONSORING AGENCY CODE EPA/600/08	
15. SUPPLEMENTARY NOTES Work performed under Cooperative Agreement CR808536-01-0 with Harvard School of Public Health		
16. ABSTRACT The health effects of pollutants are often modeled by a dose-response relationship where the dose may be the mass of the pollutant delivered to a target body compartment. Since the dosage cannot be measured, the subject's personal exposure or the ambient exposure is often used as surrogates for the dosage. As shown by derivations and simulations, the usage of personal exposure as a surrogate for dosage leads to an unbiased relationship. Since the use of ambient exposure as a surrogate for personal exposure can be treated as an observation of the personal exposure with measurement error, the use of ambient exposure as a surrogate for dosage leads to a biased estimate of dose-response effect. The simulations show that personal exposure as a surrogate for dosage can lead to an increased variance of the regression coefficients relating the measured responses to the scaled dosage.		
17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS Air Pollution, Dose-Response, Monte Carlo Simulation	b. IDENTIFIERS/OPEN ENDED TERMS Dose, Exposure Intake, Surrogate Effects	c. COSAT: Field Group
18. DISTRIBUTION STATEMENT Release to Public	19. SECURITY CLASS (This Report) Unclassified	21. NO OF PAGES 12
	20. SECURITY CLASS (This page) Unclassified	22. PRICE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

STATISTICAL RELATIONSHIPS
BETWEEN EXPOSURE AND DGSE

David T. Mage

U. S. Environmental Protection Agency
Environmental Monitoring Systems Laboratory
Research Triangle Park, North Carolina 27711

The health effects of pollutants are often modeled by a dose-response relationship where the dose may be the mass of the pollutant delivered to a target body compartment. Since the dosage cannot be measured, the subject's personal exposure or the ambient exposure is often used as surrogates for the dosage. As shown by derivations and simulations, the usage of personal exposure as a surrogate for dosage leads to an unbiased relationship. Since the use of ambient exposure as a surrogate for personal exposure can be treated as an observation of the personal exposure with measurement error, the use of ambient exposure as a surrogate for dosage leads to a biased estimate of dose-response effect. The simulations show that personal exposure as a surrogate for dosage can lead to an increased variance of the regression coefficients relating the measured responses to the scaled dosage.

Introduction

The health effect of a pollutant is usually modeled as a function of the dose of the pollutant using a dose-response function. In almost all air pollution health studies performed to date, the actual measure of the dose of the pollutant to the subject is replaced by the measurement of the exposure which serves as a surrogate for the dosage.¹ This can be demonstrated by considering the following situation:

Let a subject be exposed to an atmosphere containing a pollutant at concentration $X = 1 \mu\text{g}/\text{m}^3$. If the subject inhales 2 m^3 of this atmosphere in a given interval and 50% of the pollutant is deposited in the target body compartment, then the dose to the subject during this interval is $1 \mu\text{g} = (1 \mu\text{g}/\text{m}^3)(2 \text{ m}^3)(0.5)$.

Because of intersubject variabilities of ventilation rates and deposition patterns, subjects exposed to the same atmosphere may have widely different doses. If the health effect is a function of the dose delivered to the subject, then many subjects with identical exposures may have widely different responses. This may lead to misclassification of subjects since some low exposure subjects could obtain a high dose, and high exposure subjects may obtain a low dose. Furthermore, when one attempts to analyze the health effects data which may be dependent on the dosage D by use of a surrogate exposure for the dosage, the relationship obtained may be influenced adversely by either increasing the uncertainty of the result or by introducing a possible bias into the results.² This paper discusses the effect of the use of two exposure surrogates for dosage and presents the result of a simulation study to illustrate the effects.

Models for Relating Dose and Exposure to Response

Let us assume that for any person a health effect (ΔH) may be linearly related to the dose (D) of a pollutant and the relation can be modeled by equation 1,

$$\Delta H_i = A_i + B_i D_i \quad (1)$$

where ΔH_i is the change in the i^{th} subjects' health over interval Δt
 D_i is the dose of the pollutant in Δt reaching the target body compartment of the i^{th} subject
 B_i is a random response per unit dose for the i^{th} subject, and
 A_i is a random change in health of the i^{th} subject not due to pollution D_i plus a health effect measurement error.

If we measure $\{\Delta H_i, D_i\}$ for n individuals randomly sampled from a large population, we may estimate the average value of B_i for this population, by minimizing the mean square error of the prediction of ΔH_i for a given D_i , as b

$$\text{where } b = \frac{\sum (\Delta H_i - \bar{\Delta H})(D_i - \bar{D})}{\sum (D_i - \bar{D})^2} \quad (2)$$

NOTE: In this paper $\sum$ symbolizes $\sum_{i=1}^{i=n}$

and $\overline{\Delta H}$ is the average health effect for the n subjects, and $\overline{D}$ is the average dose for the n subjects.

If B_1 has a mean value B with variance σ_B^2 and A_1 has a mean value A with variance σ_A^2 , it can be shown that the expected value of b is B . The variance of b , σ_b^2 , can be shown to be,

$$\sigma_b^2 = \frac{\sigma_A^2}{\sum (D_1 - \overline{D})^2} + \frac{\sigma_B^2 \sum D_1^2 (D_1 - \overline{D})^2}{[\sum (D_1 - \overline{D})^2]^2} \quad (3)$$

Exposure as a Surrogate for Dosage:

Suppose that we measure the individual subject's personal exposure $\overline{X}_1$ instead of the dosage D_1 , as

$$\overline{X}_1 = \frac{\int_{t_1}^{t_2} X_1 dt}{\int_{t_1}^{t_2} dt} \quad (4)$$

where t_1 and t_2 are the start and end of interval Δt ,
 X_1 is the exposure at time t , $t_1 < t < t_2$,

Dosage D_1 can be related to exposure X_1 as

$$D_1 = \int_{t_1}^{t_2} v_1 \theta_1 X_1 dt \quad (5)$$

where v_1 is the subject's ventilation rate,
 θ_1 is the subject's deposition parameter

$$\text{or } D_1 = \overline{v_1 \theta_1} \overline{X}_1 \Delta t \quad (6)$$

where $\overline{v_1 \theta_1}$ is the i^{th} subject's exposure weighted average of $v_1 \theta_1$ during interval Δt , a random variable with mean $\overline{v\theta}$ and variance $\sigma_{v\theta}^2$.
 Substituting in equation (1) and letting $\Delta t =$ a unit time interval,

$$\Delta H_1 = A_1 + B_1 \overline{v_1 \theta_1} \overline{X}_1 \quad (7)$$

If v_1 and θ_1 are unmeasured, we may assume a linear exposure-response model

$$\Delta H_1 = a + b \overline{v\theta} \overline{X}_1 \quad (8)$$

The unit exposure response estimator, $b\overline{v\theta}$ that minimizes the mean square error in predicting ΔH_1 for a given $\overline{X}_1$, can be written as

$$b \overline{v\theta} = \frac{\sum (\overline{X}_1 - \overline{X})(\Delta H_1 - \overline{\Delta H})}{\sum (\overline{X}_1 - \overline{X})^2} \quad (9)$$

As from equation 2, b can be shown to be an unbiased estimator of B , with variance σ_b^2 ,

$$\text{where } \sigma_b^2 = \frac{1}{v\bar{\theta}^2} \left[\frac{\sigma_A^2}{\Sigma(\bar{X}_1 - \bar{X})^2} + \frac{\sigma_B^2 \Sigma_1^2 (\bar{X}_1 - \bar{X})^2 + B^2 \sigma_{v\theta}^2 \Sigma \bar{X}_1^2 (\bar{X}_1 - \bar{X})}{\Sigma[(\bar{X}_1 - \bar{X})^2]^2} \right] \quad (10)$$

Intake as a Surrogate for Dose

Suppose that we measure an individual's exposure X_1 and ventilation rate v_1 . The average intake of pollutant over the unit interval Δt is $(\bar{vX})_1$, where

$$(\bar{vX})_1 = \frac{\int_{t_1}^{t_2} v_1 X_1 dt}{\int_{t_1}^{t_2} dt} \quad (11)$$

If θ_1 is unmeasured, the observed health effect as a linear function of the intake is

$$\Delta H_1 = A_1 + B_1 \bar{\theta}_1 (\bar{vX})_1, \quad (12)$$

$$\text{where } \bar{\theta}_1 = \int v_1 X_1 \theta_1 dt / \int v_1 X_1 dt \quad (13)$$

The assumed linear intake-response model is

$$\Delta H_1 = a + b \bar{\theta} (\bar{vX})_1 \quad (14)$$

where $\bar{\theta}$ is the mean deposition parameter for the population.

The unit intake-response coefficient $(b\bar{\theta})$, that minimizes the mean square error in predicting ΔH_1 for a given intake $(\bar{vX})_1$ can be written as

$$b\bar{\theta} = \frac{\Sigma[(\bar{vX})_1 - \bar{vX}][\Delta H_1 - \Delta \bar{H}]}{\Sigma[(\bar{vX})_1 - \bar{vX}]^2} \quad (15)$$

In a similar manner, it can be shown that the expected value of b is B , and its variance σ_b^2 is,

$$\sigma_b^2 = \frac{\sigma_A^2}{\bar{\theta}^2 \Sigma[(\bar{vX})_1 - \bar{vX}]^2} + \frac{\sigma_B^2 \Sigma_1^2 [(\bar{vX})_1 - \bar{vX}]^2 + B^2 \sigma_{v\theta}^2 \Sigma(\bar{vX})_1^2 [(\bar{vX})_1 - \bar{vX}]^2}{\bar{\theta}^2 (\Sigma[(\bar{vX})_1 - \bar{vX}]^2)^2} \quad (16)$$

Comparing the variance of b predicted by equations (3), (10), and (16) we see that as we gain information on exposure, then on intake and finally on dose, the variance of our estimate for B decreases.

Effect of an Ambient Surrogate for Personal Exposure

If instead of measuring personal exposure $\bar{X}_1$, we use an ambient monitoring value $\bar{X}_{a1}$ as a surrogate for $\bar{X}_1$, then we can treat $\bar{X}_{a1}$ as $\bar{X}_1$ measured with error ϵx_1 . That is

$$\bar{X}_{a1} = \bar{X}_1 + \epsilon x_1 \quad (17)$$

If we model the health effect as being related to X_{a1} , then

$$\Delta H_1 = a + b' X_{a1} \quad (18)$$

where b' is the response coefficient of the i th subject to the ambient exposure $\bar{X}_{a1}$.

If we minimize the squares of the errors in predicting ΔH_1 for a given $\bar{X}_{a1}$

$$b' = \frac{\sum (\Delta H_1 - \bar{\Delta H})(\bar{X}_{a1} - \bar{X}_a)}{\sum (\bar{X}_{a1} - \bar{X}_a)^2} \quad (19)$$

where $\bar{X}_a$ is the average ambient exposure for the n subjects.

Substituting equation (17) in equation (19),

$$b' = \frac{\sum (\Delta H_1 - \bar{\Delta H})(\bar{X}_1 + \epsilon x_1 - \bar{X}_a)}{\sum [(\bar{X}_1 - \bar{X}_a) + \epsilon x_1]^2} \quad (20)$$

Noting that $\bar{X}_a = \bar{X} + \bar{\epsilon} x_1$

$$b' = \frac{\sum (\Delta H_1 - \bar{\Delta H})[(\bar{X}_1 - \bar{X}) + (\epsilon x_1 - \bar{\epsilon} x_1)]}{\sum [(\bar{X}_1 - \bar{X}) + (\epsilon x_1 - \bar{\epsilon} x_1)]^2} \quad (21)$$

Comparing (21) with equation (9) we see that b' will be a biased estimator of $b \sqrt{\theta}$ due to the presence of the finite quantity $\sum (\epsilon x_1 - \bar{\epsilon} x_1)^2$ in the denominator.

Therefore, personal exposure $\bar{X}_1$ as a surrogate for dosage $\bar{D}_1$ leads to an unbiased estimator of b but ambient exposure $\bar{X}_{a1}$ is a surrogate for personal exposure which can be treated as an exposure observation with measurement error which leads to a biased estimator.^{2,3}

Simulation of Surrogate Effects

To demonstrate the effect of choice of a surrogate for dosage the following model was constructed:

$$\Delta H_1 = A_1 + B_1 D_1 \quad (1)$$

$$D_1 = v_1 \theta_1 \bar{X}_1 \Delta t \quad (6)$$

To make the simulation realistic the following submodels were used for $\bar{X}_1$, v_1 and θ_1 . Note Δt is set here equal to 1 to simplify the scaling.

$$A_1 = 0.25 * N(0) \quad (23)$$

$$B_1 = \exp[N(0)]/1.649 \quad (24)$$

$$\bar{X}_1 = 0.12 \exp[N(0)] \quad (25)$$

$$v_1 = 10.5 \exp[0.31 N(0)][1 + 0.5 \exp[N(0)]] \quad (26)$$

$$v_{nasal} = v_1(1 - \exp[-.693(40)/v_1]) \quad (27)$$

$$v_{oral} = v_1 - v_{nasal} \quad (28)$$

$$\theta_{nasal} = 0.014 \exp[-5/v_{nasal}]\exp[0.35 N(0)] \quad (29)$$

$$\theta_{oral} = \exp[-18.4/v_{oral}]\exp[0.35 N(0)] \quad (30)$$

$$D_1 = \bar{X}_1[v_{nasal} \theta_{nasal} + v_{oral} \theta_{oral}] \quad (31)$$

The simulated situation is for a gaseous pollutant, such as SO_2 ,⁴⁻⁶ which can penetrate to the lower pulmonary region through either the oral or nasal tract. The subjects breathe oronasally with the fraction of the oral portion increasing as the ventilation increases. By this model, at $v_1 = 40$ lpm, 50% is in each of the oral and nasal tracts. The oral and nasal depositions are different and dependent on the ventilation rates. The personal exposures, depositions, and ventilations are also random variables through the use of the standard normal deviates symbolized by $N(0)$ representing random samples from a normal distribution with mean zero and unit variance. In this formulation B_1 is a random variable signifying the difference in sensitivities the subjects have to the pollutant. In this case, B_1 is scaled so that $E[B] = 1$ and A_1 is a random variable with $E[A] = 0$.

The simulation was run for 10 cohorts of 100 subjects each to simulate studies of 100 subjects whose health status and personal exposures are measured without error, their ventilation rates are measured during the study, and deposition parameters are known from independent measurements made previously.

The regressions were run on 3 models.

$$\text{Model 1} \quad \Delta H = A + \bar{B} (\theta v X)$$

$$\text{Model 2} \quad \Delta H = A + \bar{B} \theta (v X)$$

$$\text{Model 3} \quad \Delta H = A + \bar{B} \theta v (X)$$

In model 1, ΔH is regressed against deposited Dose D to obtain an estimate $\bar{B}$ for B . In model 2, ΔH is regressed against the dose inhaled (vX), the total mass entering the oronasal tract - a fraction of which is exhaled. In model 3, ΔH is regressed against the exposure X . Therefore, we expect the variance of the estimates to increase as we go from $\bar{B}$ to $\bar{B}\theta$ to $\bar{B}\theta v$ as the estimated coefficients. The results of the 3 regressions are shown in Table I.

TABLE I

RESULTS OF MODELS FOR ΔH WITH DOSE AND SURROGATES
10 Cohorts with 100 Observations Each

Model 1 $\Delta H = A + \bar{B}D$		Model 2 $\Delta H = A + \bar{B}\theta(vX)$		Model 3 $\Delta H = A + \bar{B}v\theta(X)$	
$\bar{B}$	t	$\bar{B}\theta$	t	$\bar{B}v\theta$	t
0.356	1.35	0.0084	1.15	-0.013	-0.05
0.481	2.62	0.0135	1.52	0.076	0.52
0.552	6.47	0.0145	2.46	0.132	0.89
0.607	6.85	0.0184	2.22	0.135	1.50
1.055	9.68	0.0190	1.34	0.137	1.15
1.142	5.94	0.0261	2.70	0.262	1.63
1.388	5.67	0.0263	4.15	0.306	2.29
1.701	9.19	0.0590	3.70	0.389	1.50
1.830	11.97	0.0654	4.32	0.467	1.51
1.930	9.81	0.0865	6.48	1.500	4.19
Avg $\bar{B} = 1.104$		Avg $\bar{B}\theta = 0.0335$		Avg $\bar{B}v\theta = 0.341$	
$S_B = 0.592$		$S_{B\theta} = 0.0267$		$S_{B\theta v} = 0.434$	
C.O.V. = 0.536		C.O.V. = 0.797		C.O.V. = 1.273	

$$\text{Max } \bar{B} - \text{Min } \bar{B} = 1.4$$

$$\text{Avg } \bar{B}$$

$$\text{Max } \bar{B}\theta - \text{Min } \bar{B}\theta = 2.3$$

$$\text{Avg } \bar{B}\theta$$

$$\text{Max } \bar{B}v\theta - \text{Min } \bar{B}v\theta = 4.4$$

$$\text{Avg } \bar{B}v\theta$$

Discussion of Results

These simulation results (with Δt scaled = 1) demonstrate the effect of the exposure surrogate for delivered dosage when a dose-response is occurring. In Model 1, the expected value $E[B]$ is 1 due to the lognormal distribution of B_1 given by equation 24. The average $\bar{B}$ is 1.104 and 9 of the 10 simulations give a significant result ($t > 1.65$). In Models 2 and 3 the expected values of $\bar{B}\theta$ and $\bar{B}v\theta$ are complicated functions of the oronasal flow and depositions.

The average total ventilation rate from equation 26 is about 20 lpm, and the weighted deposition fraction at 20 lpm is 0.015 by equation 31. The values of $\bar{B}\theta$ and $\bar{B}v\theta$ at 20 lpm are 0.015 and 0.3 respectively. These are not the expected values of $\bar{B}\theta$ and $\bar{B}v\theta$, but they serve to demonstrate their expected order of magnitudes.

In Model 2, the average $\bar{B}\theta$ of the 10 runs is 0.0335, and 7 of the 10 values are significant ($t > 1.65$). In Model 3, the average value of $\bar{B}v\theta$ is 0.341, and only 2 of the 10 values are significant. The sample variance of the estimators varies with the models as shown by the coefficient of variation (C.O.V.) varying from 0.54 in Model 1, to 0.80 in Model 2, to 1.27 in Model 3. The (Maximum-Minimum)/Mean also increases from 1.4 in Model 1 to 2.3 in Model 2 to 4.4 in Model 3.

These results demonstrate the expected loss of precision in estimating the parameter $\bar{B}$ when exposure X and intake vX are surrogates for the delivered dose $vX\theta$. As opposed to the use of the ambient exposure $\bar{X}_a$ as a surrogate for dosage which produces a biased (scaled) estimate of $\bar{B}$, the use of the personal exposure $\bar{X}_1$ for the subjects leads to an unbiased (scaled) estimator of $\bar{B}$. It is, therefore, recommended that studies of subject cohorts measure their individual personal exposures and make an effort to characterize the individual subject's physiological functions that relate to ventilation and its variability with physical activity, as well as the oronasal partition. The individual deposition fractions ($\bar{\theta}_1$) may be difficult to obtain, but their variation could be modeled if the ventilation and oronasal pattern of breathing is known.

REFERENCES

1. D. T. Mage, "Public health aspects of air quality surveillance," *Public Health Reviews* 11, 5-54 (1983).
2. J. H. Pickles, "Air pollution estimation error and what it does to epidemiological analysis," *Atmos. Environ.* 9, 2241-2245 (1982).
3. R. B. Davies and B. Hutton, "The effect of errors in the independent variables in linear regression," *Biometrika* 62, 383-391 (1975).
4. V. Niinima, P. Cole, S. Mintz, and R.J. Shepard, "Oronasal distribution of respiratory air flows," *Respir. Physics* 43, 69-75 (1981).
5. G. N. Melville, "Changes in specific airway conductance in healthy volunteers following nasal and oral inhalation of SO₂," *W. I. Med. J.* 19, 231-235 (1970).
6. F. E. Speizer and N. R. Frank, "The uptake and release of SO₂ by the human nose," *Arch. Environ. Health*, 12, 725-728 (1966).